Phenotypic Diversity of Modern Chinese and North American Soybean Cultivars

Zhanglin Cui, Thomas E. Carter, Jr.,* Joseph W. Burton, and Randy Wells

ABSTRACT

Chinese and North American (NA) soybean breeding programs have a 70-yr history of genetic progress in relative isolation from each other. Because both programs rest upon a genetic base that is primarily Chinese in origin, the actual genetic distinctness of Chinese and NA breeding is not clear. The objectives of this study were to (i) develop a phenotypic similarity (PS) index for a large group of Chinese and NA cultivars, on the basis of biochemical, morphological, and agronomic traits, (ii) compare Chinese and NA cultivars for PS through cluster analysis, and (iii) use results to develop guidelines for management of the contrasting Chinese and NA breeding programs as reservoirs of diversity. Chinese (47) and NA (25) cultivars were evaluated for 25 traits in growth chambers. Traits pleiotropic to maturity were avoided. Significant ($P < 0.05$) differences between Chinese and NA cultivars were noted for leaf and seed traits. Multivariate analysis captured 79% of the total genotypic variation among the 72 cultivars and was used to develop PS estimates. Cluster analysis of PS showed a much greater phenotypic diversity among Chinese than among NA cultivars and a striking distinctness between the two groups. The contrasting nature of Chinese and NA cultivars in this study is theorized to reflect that (i) the NA cultivars may trace to a subset of the Chinese cultivar genetic base, and/or (ii) Chinese and NA cultivars may have diverged phenotypically via breeder selection pressure. Cluster results here, based on PS, agreed roughly with previous cluster analyses, which were derived from pedigree analysis. The physical distinctness of NA and Chinese cultivars shows that introgression of Chinese cultivars into NA breeding should broaden NA germplasm's agronomic, morphological, and biochemical diversity. Introgression may be accomplished most effectively by avoiding matings of Chinese and NA cultivars from the same phenotypic cluster.

Genetic diversity is important to applied crop breeding, because diversity may reduce vulnerability to pests and, at the same time, accelerate breeding progress for an agronomic trait such as yield. In China, the continuing use of diverse landraces and exotic modern cultivars as parental stock in recent decades has been associated with improved yield and pest resistance in soybean [*Glycine max* (L.) Merr.] (Gai, 1997; Cui et al., 1998, 2000a,b). In North America, breeding progress has also been great (Specht and Williams, 1984; Specht et al., 1999). However, progress in North America has rested upon a narrow genetic base: 14 dominant ancestors introduced to North America prior to 1930 (Gizlice et al., 1993, 1994). The rate and duration of future yield gains in NA soybean breeding is a topic of current debate (Sneller, 1994; Zhou et al., 1998; Sneller, 1999; Fehr, 1999).

Carter et al. (2000) suggested that Chinese cultivars may be an important reservoir of diversity with which to expand the genetic base of NA breeding. In that regard, China has released 651 cultivars from 1923 to 1995 (Cui et al., 1999). Through pedigree analysis, Cui et al. (2000a,b) established that Chinese cultivars were derived from a far greater number of ancestors than were NA cultivars (339 vs. 80 ancestors) and that the two regions had few identifiable ancestors in common. Carter et al. (2000) also demonstrated that several Chinese cultivars performed well agronomically in North America.

In China, soybean breeders have recognized the potential importance of China and NA as mutual reservoirs of diversity. They have released 67 cultivars since 1974 with at least 25% of the pedigree derived from modern NA cultivars (Cui et al., 1999, 2000a). Many of these cultivars are high yielding, a finding that is consistent with the general notion that modern Chinese and NA cultivars may contrast in a beneficial way.

Despite important signals pointing to the desirability of exotic cultivars in Chinese and NA soybean breeding, the true utility of cross breeding these two cultivar pools, especially in terms of NA yield improvement, remains an open question. No NA cultivar has been developed from modern Chinese cultivars. No unique yield genes have been identified in these two breeding pools.

The uniqueness of Chinese and NA breeding pools have been elucidated to date primarily through pedigree analysis. The value of cross-breeding Chinese with NA cultivars will be greater than otherwise if the two groups of cultivars are distinct not only in terms of published pedigree, but in the underlying genetics as well. A limiting factor in the pedigree analyses of diversity for Chinese and NA cultivars is that, although the genetic bases of North America and China contrast in size and specific members, most members of these two bases trace their origins to China. Thus, it is possible that Chinese and NA cultivars may be derived from a common preexisting genetic base despite the lack of supporting pedigree evidence. In that regard, as many as 20 000 landraces were available in China by 1900 for use as parents in modern NA and Chinese breeding. There is no clear historical record which relates the selection of founding stock for Chinese and NA breeding. Relatively few of these landraces were used as breeding stock in North America or China, while many of these landraces are preserved in China, but the genetic structure of this large collection is not well understood (Chang et al., 1992, 1999).

One approach to elucidate underlying genetic differences between NA and Chinese breeding pools is to estimate genetic distinctness between them on the basis

Z. Cui and R. Wells, Dep. of Crop Science, North Carolina State Univ., Raleigh, NC 27695-7631; T.E. Carter, Jr. and J.W. Burton, USDA-ARS and Dep. of Crop Science, North Carolina State Univ., Raleigh, NC 27695-7631. Received 21 Dec. 2000. *Corresponding author (tommy_carter@ncsu.edu).

Published in Crop Sci. 41:1954–1967 (2001).

Abbreviations: CP, coefficient of parentage; NA, North American; PS, phenotypic similarity; N, northern NA; S, southern NA; NEC, northeastern China; NC, northern China, i.e., Huanghe, Huaihe, and Haihe valleys; SC, southern China; SEPEL, Southeastern Plant Environment Laboratories.

of measures other than pedigree. Molecular markers have been used to estimate genetic differences in germplasm accessions of soybean and other crops (Li et al., 2001; Thompson et al., 1997; Autrique et al., 1996; Johns et al., 1997). Preliminary analysis of simple sequence repeat (SSR) markers indicates that Chinese and NA soybean cultivars form fairly distinct breeding pools (Carter et al., 2000). Li et al. (2001) showed that ancestors of NA and Chinese breeding differ in terms of random amplified polymorphic DNA (RAPD) markers.

Phenotypic differences may also elucidate genetic differences. In the context of Chinese and NA breeding, we suggest that distinctions between the two pools, exemplified by pedigree and DNA marker analysis, should also be expressed phenotypically. Theoretically, phenotypic diversity should approximate genetic diversity. As the number of phenotypic traits increases in a comparison of breeding pools, the number of genes involved in the control of phenotypic traits should increase accordingly and, thereby, improve the utility of phenotypic diversity in predicting genotypic diversity. Employing this concept, van Beuningen and Busch (1997), Johns et al. (1997), and Autrique et al. (1996) used morphological, developmental, and physiological traits to create distance measures and examine genetic diversity in large collections of crop genotypes. Grafius et al. (1976) and Grafius (1978) were among the first to apply this concept to practical breeding by employing cultivar differences in morphological traits to select genetically diverse breeding pairs. In their work, the limited backcrossing of contrasting quantitative seed traits (presumably controlled by an array of genes on multiple linkage groups) from one cultivar into another tended to improve seed yield.

Recently, many modern Chinese cultivars were evaluated extensively in a series of field tests in North America for yield and other agronomic traits (Carter et al., 2000). However, we could find no comparison of genotypic diversity in NA and Chinese cultivars on the basis of a wide array of phenotypic traits. The objectives of this study were to (i) quantify diversity for 72 representative modern NA and Chinese cultivars by a phenotypic similarity (PS) index based on biochemical, morphological, and agronomic traits, (ii) compare Chinese and NA cultivars for PS through cluster analysis, and determine the relation between PS and previously reported coefficient of parentage (CP) estimates, and (iii) incorporate results into guidelines for management of the contrasting breeding programs as reservoirs of diversity.

MATERIALS AND METHODS

Forty-seven soybean cultivars from China and 25 from North America were selected for this study. They represented a broad range in genetic diversity in their respective countries on the basis of pedigree, maturity, and geographical origin (Table 1). The CP relationships among Chinese cultivars and their pedigrees were obtained from Cui et al. (1998, 1999, 2000a). The CP relationships among NA cultivars were obtained from Carter et al. (1993). Five of the 47 Chinese cultivars were derived, in part, from NA cultivars (Table 1). The proportion of NA pedigree in these Chinese cultivars was estimated by Cui et al. (1998). The relative yield of Chinese soybean cultivars in comparison with standard NA types was obtained from Carter et al. (2000). Seed were obtained from the USDA-ARS Soybean Germplasm Collection at Urbana, IL.

The experiment was conducted in 1997 in the phytotron facilities of the Southeastern Plant Environment Laboratories (SEPEL) at North Carolina State University. The 72 entries were evaluated in six 2.42- by 3.63- by 2.12-m chambers employing a randomized complete block design with chamber designated as a block and pot as the experimental unit. The planting dates for the six chambers were 10 June, 27 June, 27 June, 18 September, 2 October, and 2 October. Six healthy seeds were planted directly into each 25.6-cm-diam pot. Eight days after planting (DAP), the pots were thinned to three plants. At 12 DAP, one additional plant was removed, leaving the two most vigorous and uniform plants for the remainder of the study. The photoperiod was 19 h (from 0800–0300 h) at planting to delay flowering and sustain vegetative growth. At 14 DAP, photoperiod was reduced to 12 h (0800–2000 h) to induce prompt flowering of all genotypes. The first genotype flowered 18 d after photoperiod was reduced. Temperatures were maintained at 26 and 22°C for light and dark periods, respectively. Standard SEPEL protocols were followed for watering, fertilization, substrate preparation, and light intensity (Downs and Bonaminio, 1976). No *Rhizobium* inoculation was applied. The above methodology was similar to that employed by Gizlice et al. (1993).

Traits Evaluated

The traits measured before flower initiation were as follows. (i) Leaf length and leaf width: the most recently fully developed leaves (2nd and 3rd trifoliate leaves) and the petioles were removed from each plant at 28 DAP. Leaflets were measured for length and width. (ii) Leaf Ratio: ratio of leaf length to width. Area per leaf: the area of a leaf determined using a leaf area meter for those leaflets employed in length measurements. (iii) Petiole length: petioles of leaves used in leaf area measurements were excised and measured. (iv) Dry weight per leaf: leaf samples were dried at 60°C for 24 h and weighed. (v) Specific leaf weight: the ratio of dry weight per leaf to area per leaf. (vi) Leaf nitrogen content: leaf samples from the leaf area measurement were dried and ground. The N content of samples was determined by the Kjeldahl method (Nelson and Sommers, 1973; Glowa, 1974). (vii) Chlorophyll a, b, and total content: three leaf disks 6.5 mm in diameter were punched from each leaflet of the 4th trifoliate. The chlorophyll content of the samples was determined by means of the technique described by Moran and Porath (1980) and Moran (1982). (viii) Chlorophyll ratio: ratio of chlorophyll a to b. (ix) Vegetative plant height and node number: height and number of nodes recorded at 28 d after planting. (x) Vegetative internode length: the ratio of plant height to number of nodes at 28 d after planting.

The traits measured after maturity were as follows. (i) Stem diameter: main stem diameter measured between the first and second nodes using a caliper. (ii) Internode length: the ratio of plant height to number of nodes at harvest. Number of seeds per pod: ratio of number of seeds per plant to number of pods per plant. (iii) Hundred-seed weight: weight/100 seed. (iv) Seed protein content and seed oil content: seed protein and oil contents were determined by near infra-red (NIR) analysis at USDA Northern Regional Research Center in Peoria, IL. All seeds were yellow, which is a prerequisite for the NIR analysis. (v) Palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3)

Table 1. Seventy-two soybean cultivars employed in a phytotron study, their state or province of origin, growing region, phenotypic similarity (PS)-based cluster, coefficient of parentage (CP)-based cluster, percentage of North American (NA) pedigree, relative yield, and two-dimension coordinates obtained from multidimensional scaling (MDS) of PS for Fig. 1.

	Cultivar			Descriptive information								MDS coordinates for PS	
Code	U.S. PI designation†	ID‡	Name	State or province of origin	Growing region§	CP-based cluster¶	NA pedigree#	Yield relative to NA cultivars††	U.S. MG‡‡	Year of release	PS-based cluster§§	Dimension 1	Dimension 2
							%	%					
					North American cultivars								
1	548659	148	Braxton	Florida	S	6	100		7	1979	G	−0.67	−0.04
2	548519	209	BSR 101	Iowa	N	8	100		1	1985	F	0.22	0.20
3	548512	150	Century	Indiana	N	4	100		2	1979	E	−0.09	0.19
4	508083	223	Dassell	Minnesota	N	9	100		0	1986	E	−0.04	0.28
5	592756		Dillon	South Carolina	S		100		6	1996	F	0.22	0.12
6	513382	241	Glenwood	Minnesota	N	3	100		0	1987	E	−0.01	0.10
7	518664	244	Hutcheson	Virginia	S	7	100		5	1987	F	0.28	0.18
8	508267		Johnston	North Carolina	S		100		8	1983	E	0.06	0.06
9	586981		KS 4694	Kansas	N		100		4	1997	A	0.24	0.02
10	562373		Lambert	Minnesota	N		100		0	1992	E	0.09	0.29
11	593258		Macon	Illinois	N		100		3	1996	C	−0.15	−0.10
12	559932		Manokin	Maryland	N		100		4	1991	E	0.09	0.06
13	548642	225	Maple Donovan	Ontario	N	3	100		0	1986	F	0.23	0.24
14	548582	137	McCall	Minnesota	N	9	100		−1	1978	E	0.05	0.24
15	510670	226	Morgan	Maryland	N	1	100		4	1986	E	−0.05	0.00
16	542404	195	Ozzie	Minnesota	N	5	100		0	1983	C	−0.18	−0.04
17	563374		Parker	Minnesota	N		100		1	1992	E	0.02	0.13
18	548523	156	Pella	Iowa	N	8	100		3	1979	E	−0.12	0.36
19	548520	215	Preston	Iowa	N	1	100		2	1985	E	0.03	−0.04
20	508084	231	Sibley	Minnesota	N	5	100		1	1986	E	−0.09	0.11
21	531068	255	Stonewall	Alabama	S	2	100		7	1988	E	−0.01	0.01
22	542768		Sturdy	Minnesota	N		100		1	1990	E	0.02	0.22
23	548991	234	TN 5-85	Tennessee	S	2	100		5	1986	E	−0.01	0.04
24	548524	158	Weber	Iowa	N	4	100		1	1979	E	0.02	0.33
25	508266	207	Young	North Carolina	S	7	100		6	1984	F	0.29	0.15
					Chinese cultivars								
26	467317	C323	De Dou 1 Hao	Jilin	NEC	D	0	81	1	1985	A	0.16	−0.14
27	467323A	C383	Jiu Nong 13	Jilin	NEC		0	82	0	1981	C	−0.02	−0.15
28	503334	C472	Dan Dou 5 Hao	Liaoning	NEC		0	61	3	1981	E	−0.17	0.12
29	503336	C143	Dong Nong 37	Heilongjiang	NEC	D	0	75	−1(0)	1984	A	0.17	−0.35
30	503340	C398	Tong Nong 9 Hao	Jilin	NEC		0	72	2	1987	C	−0.15	−0.23
31	511866	C259	Nen Feng 9 Hao	Heilongjiang	NEC	C	0	90	0	1980	C	−0.11	−0.07
32	511867	C260	Nen Feng 10 Hao	Heilongjiang	NEC	C	0	94	0	1981	C	0.05	−0.07
33	518706A	C214	Hei Nong 29	Heilongjiang	NEC	B	0	67	0	1986	C	−0.07	−0.20
34	518709	C350	Ji Lin 18	Jilin	NEC		0	72	1(2)	1982	A	0.37	−0.08
35	518710	C352	Ji Lin 20	Jilin	NEC		0	76	1	1985	A	0.24	−0.31
36	518711	C353	Ji Lin 21	Jilin	NEC		0	75	2	1988	E	0.09	0.17
37	518712	C494	Kai Yu 8 Hao	Liaoning	NEC	H	0	72	2	1980	E	0.16	0.04
38	518714	C503	Liao Nong 2 Hao	Liaoning	NEC	H	0	75	2	1983	A	0.22	−0.07
39	518718A	C545	Lu Dou 4 Hao	Shandong	NC		0	65	2(3)	1985	A	0.32	−0.01
40	518719	C548	Lu Dou 7 Hao	Sandong	NC		0	66	4	1987	E	−0.06	0.14
41	532459	C423	Huai Dou 1 Hao	Jiangsu	SC		0	59	4	1983	E	−0.12	0.24
42	549076A	C226	Hong Feng 3 Hao	Heilongjiang	NEC	D	0	75	−1(0)	1981	A	0.13	−0.27
43	549077	C234	Jiu Feng 1 Hao	Heilongjiang	NEC		0	74	0	1983	A	0.00	−0.27
44	549078	C235	Jiu Feng 2 Hao	Heilongjiang	NEC		0	70	−2(−1)	1984	A	0.12	−0.16
45	592920	C178	He Feng 33	Heilongjiang	NEC	J	0	81	0	1992	B	−0.08	−0.46
46	592921	C222	Hei Nong 37	Heilongjiang	NEC		0	91	1	1992	C	−0.24	−0.08
47	592923	C247	Ken Nong 2 Hao	Heilongjiang	NEC	E	0	83	0	1988	B	−0.18	−0.45
48	592925	C314	Bai Nong 1 Hao	Jilin	NEC	G	0	68	0	1981	C	0.03	−0.16
49	592926	C397	Tong Nong 8 Hao	Jilin	NEC	F	0	85	1	1982	E	−0.22	0.08
50	592927	C619	Chen Dou 4 Hao	Sichuan	SC		0	54	2	1989	A	0.40	0.07
51	592928	C628	Gong Dou 4 Hao	Sichuan	SC		25	50	2(3)	1992	D	−0.28	0.18
52	592929	C421	Guan Dou 1 Hao	Jiangsu	SC		0	50	4	1985	C	−0.22	−0.25
53	592932	C033	Zao Shu 9 Hao	Beijing	NC		0	47	2	1983	D	−0.40	0.16
54	592933	C034	Zao Shu 14	Beijing	NC		0	61	2	1987	C	−0.23	−0.28
55	592934	C649	Zhe Chun 2 Hao	Zhejiang	SC		0	61	2	1987	A	0.24	−0.16
56	592936	C085	Ji Dou 7 Hao	Hebei	NC		50	83	2	1992	C	−0.23	−0.19
57	592937	C604	Jin Dou 14	Shanxi	NC		0	65	4	1991	D	−0.31	0.33
58	592938	C605	Jin Dou 15	Shanxi	NC	L	0	75	2(1)	1991	E	−0.17	0.10
59	592939	C606	Jin Dou 16	Shanxi	NC	L	0	68	4	1991	E	−0.18	0.17
60	592941	C502	Liao Dou 10 Hao	Liaoning	NEC	J	25	70	2(3)	1992	B	0.06	−0.46
61	592942	C533	7605	Shandong	NC		0	68	3	1986	A	0.30	−0.24
62	592944	C514	Tie Feng 22	Liaoning	NEC	C	0	74	2(3)	1986	A	0.28	0.27
63	592945	C038	Zhong Huang 1 Hao	Beijing	NC	K	0	58	3	1989	E	−0.06	0.29
64	592946	C082	Ji Dou 4 Hao	Hebei	NC		50	71	4	1984	A	0.06	−0.28
65	592947	C612	Jin Yi 9 Hao	Shanxi	NC	K	0	82	4	1989	E	0.10	0.19
66	592948	C613	Jin Yi 10 Hao	Shanxi	NC	K	0	58	3	1988	E	0.06	0.11
67	592949	C111	Yu Dou 8 Hao	Henan	NC	O	0	57	4	1988	A	0.22	0.00
68	592950	C113	Yu Dou 11	Henan	NC	O	0	65	4	1992	C	−0.25	−0.04
69	592951	C120	Zheng 133	Henan	NC		0	37	4	1990	C	−0.31	0.01

Continued next page.

Table 1. Continued.

Cultivar				Descriptive information								MDS coordinates for PS	
Code	U.S. PI designation†	ID‡	Name	State or province of origin	Growing region§	CP-based cluster¶	NA pedigree#	Yield relative to NA cultivars††	U.S. MG‡‡	Year of release	PS-based cluster§§	Dimension 1	Dimension 2
							%	%					
						Chinese cultivars							
70	592952	C121	Zheng 77249	Henan	NC	O	0	63	3	1983	B	−0.38	−0.20
71	592953	C295	Zhong Dou 19	Hubei	NC	Q	0	54	4	1987	C	−0.21	−0.01
72	592954	C438	Nin Zhen 1 Hao	Jiangsu	SC	Q	50	65	2	1984	A	0.32	−0.12

† Plant introduction number.
‡ Identification numbers were consistent with Carter et al. (1993) or Cui et al. (1998, 1999).
§ N = northern NA, S = southern United States, NEC = northeastern China, NC = northern China, SC = southern China.
¶ Coefficient of parentage (CP) based clusters were obtained from Gizlice et al. (1996) for NA cultivars and from Cui et al. (2000b) for Chinese cultivars. Both were derived from non-hiarachical FASTCLUS procedure of SAS.
# Percent of genes traced to NA soybean lines.
†† Percent of yield as compared with NA check cultivars in U.S. field test (Carter et al., 2000).
‡‡ MG = maturity group, data from Germplasm Resources Information Network (GRIN). Data in parentheses were from field evaluations in multiple locations and two years and represented a discrepancy with GRIN data (Carter et al., 2000). Maturity designations were presented as Arabic rather than Roman numerals for ease of presentation.
§§ Phenotypic-based phenotypic similarity (PS) clusters were derived from Ward's minimum variance cluster analysis.

contents in seed oil were determined by gas chromatography as described by Carver et al. (1984).

Days to flowering and days to maturity from planting, flower, and pubescence color were also recorded. Maturity Group was determined from a field study (Carter et al., 2000). Traits were measured for all six chambers with the following exceptions. Protein and oil content, and fatty acid composition of oil were determined for four chambers. Chlorophyll, plant height, and leaf N were determined for five chambers.

Statistical Analysis

Analysis of variance and least squares genotypic means were computed for all traits by the GLM procedure of SAS (SAS Institute, 1985b). Normality of genotypic means was tested by the UNIVARIATE procedure of SAS (SAS Institute, 1985a). The Shapiro-Wilk statistic, W, indicated that genotypic means approximated normal distributions except for the trait, leaf ratio, which was dropped from further analysis.

Principal component analysis was performed on the standardized least squares genotypic means to calculate eigenvalues, eigenvectors, and principal component scores calculated by the PRINCOMP procedure of SAS (SAS Institute, 1985b). Pairwise genetic distances among the 72 genotypes were computed as follows:

$$D_{ij} = \left[\sum_{k=1}^{q} (y_{ik} - y_{jk})^2/\lambda_k\right]^{\frac{1}{2}}$$

Where, D_{ij} = genetic distance between ith and jth genotypes; y_{ik} and y_{jk} are the kth principal component scores for Genotypes i and j; λ_k is the kth largest eigenvalue; λ_q is the smallest eigenvalue that is >1.0 (Goodman, 1972; Gizlice et al., 1993).

Phenotypic similarity was derived from genetic distance by the formula:

$$PS_{ij} = 1 - (D_{ij}/D_{\max})$$

Where, PS_{ij} = phenotypic similarity between ith and jth genotypes; D_{ij} = genetic distance between ith and jth genotypes; and $D_{\max}$ is the largest genetic distance.

Ward's minimum variance method and eleven other clustering techniques were applied to PS scores, employing the CLUSTER procedure of SAS (SAS Institute, 1985b). A separate additional nonhierarchical clustering approach was also used, where multidimensional scaling (MDS) was applied to produce Euclidean coordinates (71 dimensions) based on the 72 by 72 PS matrix (SAS Institute, 1992; Gizlice et al., 1996). The nonhierarchical cluster analysis FASTCLUS procedure was then applied to MDS-derived 71-dimensional Euclidean coordinates (SAS Institute, 1985b). For each FASTCLUS analysis, it was necessary to specify the number of clusters desired prior to conducting the analysis. To find an optimum analysis, we performed 24 separate FASTCLUS analyses, specifying 2 to 25 clusters. The MEANS procedure of SAS was used to calculate mean PS within and among clusters for each analysis (SAS Institute, 1985a).

To determine the importance of a discriminator, such as cluster, maturity group, geographic origin, and yield level in explaining PS, we computed the percentage of variation accounted for by the discriminator in the PS matrix. An arithmetic shortcut described by Gizlice et al. (1996) was adopted to compute the coefficient of determination (R^2) by means of PROC GLM (SAS Institute, 1985b). Correlation analysis was applied between CP and the phenotypic similarity measures derived from the phytotron data by the CORR procedure in SAS (SAS Institute, 1985b).

The PS matrix for the 72 cultivars was subjected to the multidimensional scaling (MDS) procedure of SAS with options LEVEL = ABSOLUTE, DIMENSION = 2, and SIMILAR = 1 (SAS Institute, 1992). Two-dimensional graphs were constructed from MDS-derived coordinates (Table 1) by the PLOT procedure (SAS Institute, 1985a). The graphs depicted approximate genetic distances among entries conveniently in the same units as the original PS values (Gizlice et al., 1996).

RESULTS AND DISCUSSION

Selection of Cultivars for Study

The 47 Chinese cultivars in this study were chosen because they are representative of modern Chinese soybean breeding and were released during 1980 to 1992. The cultivars were released from a total of 12 soybean-producing provinces encompassing all three of the major growing regions of China: northeastern China (NEC), northern China (NC), and southern China (SC) (Table 1). Cui et al. (2000b) identified 20 major clusters of Chinese cultivars on the basis of CP. The Chinese cultivars studied here represented 12 of the 20. The maturity of the Chinese cultivars in this study ranged

Table 2. Genotypic-based mean, variance, and range for 25 traits of soybean cultivars from northeastern China (NEC), northern China (NC), southern China (SC), all of China, northern or Midwestern part of North America (N), southern USA (S), and all of North America (NA).

		Leaf trait										
Region or country	Statistics	Leaf length	Leaf width	Area per leaf	Dry weight per leaf	Specific leaf weight	Petiole length	Leaf nitrogen content	Chlorophyll a content	Chlorophyll b content	Chlorophyll a+b content	Chlorophyll a/b ratio
		mm	mm	mm^2	mg	g m^{-2}	Mm	g kg^{-1}	g kg^{-1}	g kg^{-1}	g kg^{-1}	ratio
NEC	Mean	125.7	64.0	17 637.0	412.3	23.5	101.4	55.2	88.2	27.9	116.3	3.1
	Variance	73.0	127.0	7 305 482.8	4 441.8	2.8	475.5	9.4	120.3	8.2	186.8	0.0
	Range†	35.0	35.0	9 088.0	240.0	5.9	77.0	14.8	42.6	11.6	54.2	0.4
NC	Mean	118.9	73.2	19 348.6	472.2	22.4	124.7	54.5	90.2	28.5	118.7	3.2
	Variance	329.8	221.4	25 876 131.6	14 810.0	5.4	828.3	12.8	100.4	7.3	160.3	0.0
	Range	168.0	101.0	31 063.0	760.0	28.8	220.0	64.7	125.1	38.1	163.2	3.3
SC	Mean	121.0	77.8	20 718.0	431.5	21.4	120.3	51.6	83.9	27.2	111.1	3.1
	Variance	204.8	48.6	12 488 084.0	5 783.5	2.1	265.1	1.2	49.2	4.8	84.0	0.0
	Range	40.0	18.0	10 229.0	207.0	4.4	43.0	3.1	17.7	5.8	23.6	0.2
China	Mean	122.4	69.5	18 722.3	420.7	22.8	113.2	54.5	88.5	28.1	116.6	3.2
	Variance	195.2	179.3	16 062 771.2	8 518.1	4.2	697.6	10.8	103.6	7.3	162.9	0.0
	Range	61.0	45.0	16 887.0	403.0	9.3	111.0	19.6	42.6	11.8	54.2	0.4
N	Mean	112.8	75.9	18 958.9	433.6	22.8	116.5	57.2	102.6	31.5	134.0	3.3
	Variance	57.8	31.4	5 986 040.2	4 236.0	2.5	238.6	5.5	60.7	4.6	97.9	0.0
	Range	21.0	17.0	7 544.0	219.0	6.2	53.0	8.0	29.9	7.9	38.0	0.3
S	Mean	112.5	70.0	17 450.5	390.3	22.8	116.8	54.3	93.1	28.7	122.3	3.2
	Variance	325.4	94.6	21 727 033.4	5 347.1	16.5	614.8	10.6	33.0	1.9	36.5	0.0
	Range	53.0	27.0	13 823.0	217.0	12.4	71.0	10.9	17.2	3.5	16.2	0.4
NA	Mean	112.7	74.0	18 476.2	419.7	22.8	116.6	56.3	99.5	30.6	130.3	3.3
	Variance	133.5	56.3	10 843 460.8	4 809.3	6.4	338.4	8.7	70.5	5.5	107.3	0.0
	Range	53.0	32.0	14 097.0	251.0	13.1	71.0	12.1	31.4	8.4	38.0	0.4
Overall	Mean	119.0	71.0	18 636.8	420.4	22.8	114.4	55.1	92.3	28.9	121.4	3.2
	Variance	193.0	140.0	14 086 179.4	7 144.7	4.9	568.9	10.7	119.0	8.1	184.6	0.0
	Range	74.0	45.0	19 570.0	403.0	14.7	111.0	19.6	44.4	11.8	56.1	0.5
LSD 0.05†		6.6	5.8	1 870.4	42.2	1.1	11.9	1.6	4.8	1.3	5.9	~0.0

		Stem trait					Seed trait								
Region or country	Statistics	Vegetative plant height	Vegetative node number	Vegetative internode length	Internode length	Stem diameter	Seeds per pod	100-seed weight	Seed protein content	Seed oil content	Palmitic acid	Stearic acid	Oleic acid	Linoleic acid	Linolenic acid
		mm	no.	mm	mm	mm	no.	g	g kg^{-1}	g kg^{-1}	mole %	mole %	mole %	mole %	mole %
NEC	Mean	825.4	10.5	77.9	90.8	7.5	2.0	23.0	427.1	201.0	12.2	3.4	25.3	50.1	9.1
	Variance	24 633.2	0.2	178.7	179.7	0.9	0.0	12.6	178.9	44.7	0.6	0.1	64.8	38.0	2.2
	Range	513.0	1.9	41.0	50.0	3.6	0.7	14.2	54.0	22.0	2.7	1.5	33.7	27.3	4.8
NC	Mean	739.7	10.2	72.2	91.2	7.9	1.9	23.8	441.5	190.4	12.2	3.0	24.2	51.9	8.7
	Variance	17 153.6	0.2	123.1	152.6	0.7	0.1	11.7	306.6	217.7	0.4	0.1	39.9	22.8	1.7
	Range	1 026.0	10.6	95.0	109.0	9.3	2.0	24.9	488.0	165.0	14.5	4.0	21.9	71.4	14.5
SC	Mean	841.5	10.8	77.8	89.7	7.8	1.9	21.3	454.8	171.8	12.0	2.9	26.5	49.1	9.4
	Variance	12 054.7	0.3	91.8	124.3	1.5	0.1	18.3	173.4	51.8	0.8	0.1	34.9	21.0	1.1
	Range	293.0	1.3	25.0	25.0	2.9	0.6	11.2	37.0	21.0	2.5	0.7	16.3	12.6	3.0
China	Mean	792.8	10.4	75.6	90.8	7.7	1.9	23.1	436.5	193.0	12.2	3.2	25.0	50.7	9.0
	Variance	21 250.0	0.3	147.8	155.5	0.9	0.0	13.0	362.1	203.4	0.5	0.2	49.6	29.6	1.9
	Range	609.0	2.2	48.0	56.0	3.7	0.9	16.1	94.0	59.0	3.4	1.9	33.7	28.0	5.6
N	Mean	772.6	10.2	75.2	98.4	7.8	2.1	21.2	417.0	205.4	12.0	3.5	20.0	55.6	9.0
	Variance	10 451.6	0.2	73.4	136.4	0.2	0.0	7.1	118.1	52.9	0.2	0.1	4.0	2.2	0.6
	Range	329.0	1.2	27.0	43.0	1.6	0.4	9.4	45.0	33.0	1.7	1.1	7.0	4.7	2.8
S	Mean	801.8	10.5	75.9	93.1	7.4	2.0	21.1	421.8	205.8	12.0	3.1	22.0	53.8	9.1
	Variance	3 497.6	0.3	57.3	90.1	0.8	0.0	7.9	115.4	29.1	0.4	0.3	14.4	5.3	1.7
	Range	197.0	1.5	22.0	29.0	2.7	0.5	8.3	34.0	15.0	1.6	1.4	11.5	6.4	4.1
NA	Mean	782.0	10.3	75.4	96.7	7.6	2.0	21.2	418.5	205.5	12.0	3.4	20.7	55.0	9.1
	Variance	8 179.9	0.2	65.8	123.4	0.4	0.0	7.0	117.5	43.8	0.3	0.2	7.8	3.7	0.9
	Range	329.0	1.7	27.0	46.0	2.7	0.5	10.4	48.0	33.0	1.7	1.7	11.5	7.1	4.1
Overall	Mean	789.0	10.4	75.5	92.8	7.7	2.0	22.5	430.2	197.4	12.1	3.2	23.5	52.2	9.0
	Variance	16 559.5	0.2	118.0	150.4	0.7	0.0	11.7	348.4	182.5	0.4	0.2	39.1	24.7	1.5
	Range	609.0	2.2	48.0	57.0	3.7	0.9	16.1	102.0	74.0	3.4	2.0	34.3	28.0	5.6
LSD 0.05‡		64.1	0.3	5.4	6.0	0.4	0.1	1.6	8.3	6.0	0.3	0.2	2.9	2.3	0.6

† Difference between the largest and smallest trait values of the 72 cultivars.
‡ Least significant difference at 0.05 level for comparing the trait means of 47 Chinese cultivars and 25 NA cultivars.

from group 000 to IV, which is the maturity range for most modern soybean cultivars in China.

The 25 NA public cultivars were selected as representative of NA breeding. Seventeen were chosen to represent the nine major clusters of NA soybean cultivars identified by Gizlice et al. (1996), on the basis of pedigree analysis, and were released from 1978 to 1988. The remaining eight NA cultivars, released from 1989 to

1997, were selected as representative of more recent cultivars in both northern and southern USA. The only exception was the cultivar Johnston which was not a member of a cluster and released in 1983, but was included because of its high yield potential. Maturity ranged from 00 to VIII.

Use of Phytotron to Minimize Maturity Effects

In soybean, similarity measures based on phenotypic data are difficult to employ because of the extreme cultivar range in photoperiod sensitivity and attendant maturity. In the field, for example, genotypic differences in photoperiod sensitivity may cause one cultivar to flower just as another matures. The large pleiotropic effect of maturity on plant height, lodging, and seed yield tends to overshadow other phenotypic differences among genotypes. Thus, these traits are not readily used in the calculation of genetic distance. In addition to the clear pleiotropic effects of maturity on some plant traits, maturity also introduces a more subtle but equally important bias effect in the cultivar comparisons with respect to seed characters. A cultivar maturing more than 1 mo later than another in the same field in North America, for example, will likely experience a cooler temperature regime during pod development and, as a result, produce seed with altered size and composition (Martin et al., 1986). Thus, seed traits obtained from field studies are not well suited to studies of genetic distance when genetic differences in maturity are present. While there is not a clear solution to the problem of maturity and pleiotropism in distance estimation, Gizlice et al. (1993) demonstrated that the bias effect of maturity on seed traits is minimized when (i) phenotypic data are derived from cultivars grown in temperature-controlled growth chambers instead of field plots, and (ii) traits inherently related to maturity group, such as plant height at maturity, are deleted from analysis. Thus, growth chambers should provide an effective aid in the development of distance measures based on phenotypic traits.

The phytotron protocol employed in this study effectively minimized the maturity group effect on cultivar comparisons. Although the actual maturity groups for the 72 entries ranged from 000 to VIII (an approximate range of 9 wk in maturity in the field), the observed range in flowering and maturity dates was only 8 and 24 d, respectively. The range in flowering dates is similar to that reported for a previous phytotron-based study (Gizlice et al., 1993). Analysis of variance revealed significant ($P < 0.01$) phenotypic differences for all 25 traits studied. However, none was highly correlated with maturity group (all but one r value between ± 0.20), indicating that constant temperature regime of the phytotron was effective in removing bias effects of maturity on seed trait comparisons. Thus, the data employed in this study appeared appropriate for assessment of phenotypic diversity in Chinese and NA cultivars.

NA vs. Chinese Cultivar Differences in Phenotypic Traits

Chinese and NA cultivar groups differed in phenotypic means for 13 of 25 traits ($P < 0.05$) (Table 2). On average, Chinese cultivars exhibited longer leaves, larger seed, higher protein and lower oil content in the seed, higher oleic and lower linoleic acid content in the seed oil, and lower leaf chlorophyll and lower nitrogen contents in comparison with NA cultivars ($P < 0.05$). Chinese cultivars were also more diverse than NA cultivars for 24 of the 25 traits on the bais of range of—and variance among—genotypic means (Table 2). The NA cultivars had a broader genotypic range than Chinese cultivars only for specific leaf weight.

Comparison of growing regions within China (NEC, NC, and SC) revealed that although seed protein content is generally higher in Chinese than NA cultivars, it was the cultivars from the NC and SC regions that exhibited the highest levels of seed protein content (Table 2). The cultivars from the SC region also had the largest and thinnest leaves and lowest N and chlorophyll contents of any growing region in China or North America. Cultivars from the midwestern region of North America had larger leaves and higher N and chlorophyll content and slightly lower seed protein content than did cultivars from the southern NA region.

Phenotypic Similarity Estimates

The large genotypic variation in phenotypic traits prompted the use of multivariate analysis to identify major genotypic patterns. Principal component analysis (PCA) was used to transform cultivar least squares means (Table 3) for phenotypic traits into principal component scores. The first seven principal components (those with eigenvalues greater than one) summarized 79% of the variation in cultivar means (data not shown). The first two principal components separated Chinese and NA cultivars well, indicating a clear phenotypic distinction between the two cultivar pools in terms of leaf morphology, chlorophyll content, and seed composition.

The seven most important principal components were employed in the computation of genetic distance and subsequent phenotypic similarity (1-genetic distance) for all pair-wise combinations of the 72 cultivars. The mean PS values (and variances) within the Chinese and NA cultivar groups were 0.583 (0.012) and 0.679 (0.007), respectively, indicating that Chinese cultivars were perhaps more diverse phenotypically than the NA cultivars. The two-dimensional representation of PS relationships revealed clearly the greater phenotypic diversity in the Chinese group and the distinctness of Chinese and NA soybean cultivars (Fig. 1). The NA cultivars were located primarily in one quadrant, while Chinese cultivars were scattered broadly over the entire graph.

Cluster Analysis

Many clustering programs and algorithms are available for study of genetic diversity in crops. It is seldom clear a priori which will be best for any particular data set. To minimize the effect of clustering approach on results, we employed a number of methods. Cluster analysis was performed, initially, by the CLUSTER procedure of SAS (SAS Institute, 1985b). The PS matrix was subjected to Ward's minimum variance method and 11

Table 3. Least square means of 25 traits for 72 modern Chinese and North American (NA) soybean cultivars grown in temperature- and photoperiod-controlled growth chambers.

		Leaf trait										
Code	Name	Leaf length	Leaf width	Area per leaf	Dry weight per leaf	Specific leaf weight	Petiole length	Leaf nitrogen content	Chlorophyll a content	Chlorophyll b content	Chlorophyll a+b content	Chlorophyll a/b ratio
		mm	mm	mm^2	mg	g m^{-2}	mm	g kg^{-1}	g kg^{-1}	g kg^{-1}	g kg^{-1}	ratio
1	Braxton	77	51	8 303	321	32.5	73	48.2	91.2	26.8	118.0	3.42
2	BSR 101	116	78	19 917	414	21.0	132	57.8	111.9	34.9	146.8	3.21
3	Century	105	77	18 238	406	22.3	121	57.3	97.7	30.6	128.2	3.20
4	Dassell	102	69	15 820	376	23.8	117	58.4	102.4	31.0	133.3	3.32
5	Dillon	124	76	20 472	454	22.2	144	54.7	101.8	30.2	132.0	3.37
6	Glenwood	117	81	20 937	486	23.3	114	54.3	105.2	31.1	136.2	3.38
7	Hutcheson	118	78	20 016	408	20.4	129	55.8	98.3	30.2	128.5	3.25
8	Johnson	118	70	17 421	350	20.1	110	53.7	84.6	26.7	115.8	3.11
9	KS 4694	122	76	19 953	470	23.6	138	59.3	113.6	34.8	148.4	3.26
10	Lambert	112	72	17 566	371	21.5	109	58.1	99.3	30.5	129.9	3.26
11	Macon	120	77	20 756	523	25.5	130	53.8	116.0	35.1	151.2	3.33
12	Manokin	120	77	20 222	449	22.3	99	60.2	95.9	30.0	125.8	3.21
13	Maple Donovan	103	66	14 856	322	21.7	86	54.2	101.0	32.3	133.3	3.13
14	McCall	103	73	16 723	366	21.9	105	56.6	92.7	28.4	121.1	3.28
15	Morgan	123	82	22 028	494	22.6	124	60.3	102.8	31.8	134.6	3.23
16	Ozzie	121	80	21 032	541	25.6	125	57.3	107.4	32.1	139.6	3.33
17	Parker	104	69	15 210	363	23.9	118	57.4	101.9	32.2	134.0	3.17
18	Pella	108	80	19 476	431	22.2	103	59.3	86.1	27.2	113.2	3.18
19	Preston	120	83	22 400	522	23.4	127	59.4	109.7	32.3	142.0	3.40
20	Sibley	109	67	16 257	390	23.9	93	52.3	104.2	31.7	135.9	3.30
21	Stonewall	127	76	20 817	434	21.0	130	56.7	97.6	29.0	126.6	3.38
22	Sturdy	113	83	10 910	447	19.4	139	56.5	95.7	29.6	125.2	3.21
23	TN 5-85	130	78	22 126	507	22.9	133	59.1	89.7	27.8	117.4	3.21
24	Weber	95	60	13 034	290	22.2	87	54.2	92.7	29.2	121.9	3.16
25	Young	111	71	17 415	358	20.7	128	52.0	88.6	29.3	117.9	2.99
26	De Dou 1 Hao	130	60	16 217	390	23.9	101	52.5	94.9	30.2	125.2	3.14
27	Jiu Nong 13	130	68	19 598	450	23.1	89	56.5	92.1	28.9	121.0	3.19
28	Dan Dou 5 Hao	115	79	20 436	491	24.1	148	48.3	79.3	25.8	105.0	3.06
29	Dong Non 37	128	51	14 677	388	26.5	79	54.3	99.7	30.9	130.6	3.23
30	Tong Nong 9 Hao	122	61	16 554	421	25.3	143	57.8	88.7	27.5	116.3	3.20
31	Nen Feng 9 Hao	123	78	20 536	498	24.3	134	53.5	77.4	24.8	102.2	3.10
32	Nen Feng 10 Hao	128	80	22 050	538	24.6	133	55.5	86.7	28.1	114.8	3.08
33	Hei Nong 29	129	56	16 017	376	23.3	82	56.3	82.7	26.2	108.9	3.16
34	Ji Lin 18	122	49	12 962	314	24.3	84	55.1	89.1	27.8	116.9	3.21
35	Ji Lin 20	132	58	17 233	355	20.6	90	58.9	78.7	26.1	104.8	2.98
36	Ji Lin 21	111	71	17 610	390	22.0	100	54.9	81.6	25.7	107.3	3.17
37	Kai Yu 8 Hao	115	84	22 031	503	22.7	115	52.5	95.2	29.9	125.1	3.19
38	Liao Nong 2 Hao	122	53	13 870	298	21.5	89	54.7	71.6	23.5	95.1	3.04
39	Lu Dou 4 Hao	132	88	25 364	485	19.3	116	51.9	84.1	27.2	111.3	3.07
40	Lu Dou 7 Hao	118	77	19 853	377	19.2	99	53.4	85.3	26.6	111.9	3.21
41	Huai Dou 1 Hao	119	71	19 507	375	19.4	112	51.8	73.4	24.3	97.6	3.02
42	Hong Feng 3 Hao	125	55	15 543	374	24.1	98	56.1	107.6	32.8	140.4	3.27
43	Jiu Feng 1 Hao	136	59	18 475	451	26.1	97	54.6	114.2	35.1	149.3	3.27
44	Jiu Feng 2 Hao	128	52	15 034	358	24.5	81	53.0	100.5	30.8	131.3	3.26
45	He Feng 33	134	54	15 144	338	22.4	96	63.1	75.3	25.2	100.4	2.99
46	Hei Nong 37	119	77	20 316	497	24.4	119	54.2	89.2	27.8	117.0	3.16
47	Ken Nong 2 Hao	146	63	19 652	450	23.0	71	59.7	88.1	26.7	114.8	3.32
48	Bai Nong 1 Hao	135	54	16 480	343	20.9	85	52.7	78.7	26.0	109.5	2.93
49	Tong Nong 8 Hao	125	78	21 076	434	20.6	88	52.7	75.6	24.8	100.4	3.17
50	Chen Dou 4 Hao	129	83	22 965	492	21.6	143	52.8	78.3	25.2	103.6	3.11
51	Gong Dou 4 Hao	96	68	14 509	315	23.8	100	49.7	91.1	30.1	121.2	3.18
52	Guan Dou 1 Hao	136	86	24 738	522	21.3	134	51.4	86.1	27.4	113.5	3.16
53	Zao Shu 9 Hao	90	55	10 986	298	27.1	116	44.4	97.3	29.9	127.1	3.26
54	Zao Shu 14	150	60	18 818	431	22.9	145	53.0	93.5	28.9	122.4	3.25
55	Zhe Chun 2 Hao	116	79	20 771	437	21.3	109	52.4	83.7	27.5	111.2	3.05
56	Ji Dou 7 Hao	126	93	25 720	646	25.2	126	53.8	91.1	27.8	118.9	3.29
57	Jin Dou 14	101	69	14 793	263	17.8	125	54.2	73.4	23.3	96.7	3.12
58	Jin Dou 15	99	69	15 357	393	25.6	85	55.8	98.8	20.7	129.5	3.22
59	Jin Dou 16	98	70	15 240	318	21.3	90	54.4	83.6	29.7	111.5	3.00
60	Liao Dou 10 Hao	111	69	16 504	413	25.1	109	57.5	93.4	29.5	122.9	3.16
61	7605	131	49	14 176	322	23.1	120	54.6	100.7	30.9	131.6	3.27
62	Tie Feng 22	104	70	16 850	359	21.1	127	52.3	75.0	24.5	99.6	3.08
63	Zhong Huang 1 Hao	103	74	17 520	403	22.9	102	64.0	100.9	32.0	132.8	3.15
64	Ji Dou 4 Hao	151	54	17 802	398	22.8	110	56.3	92.8	29.6	122.4	3.15
65	Jin Yi 9 Hao	116	82	21 203	418	19.8	107	53.8	79.8	25.5	105.4	3.14
66	Jin Yi 10 Hao	104	72	16 100	375	23.4	119	55.8	99.9	31.4	131.3	3.19
67	Yu Dou 8 Hao	110	51	15 487	305	22.8	97	56.8	110.3	33.8	144.2	3.25
68	Yu Dou 11	133	94	27 555	632	23.0	159	56.6	96.1	29.3	125.4	3.29
69	Zheng 133	138	93	26 887	574	21.5	179	55.9	78.7	25.5	104.2	3.10
70	Zheng 77249	134	94	27 873	666	23.8	182	55.6	83.3	27.5	110.8	3.01
71	Zhong Dou 19	121	77	20 040	453	22.8	165	53.0	89.8	28.8	118.6	3.12
72	Nin Zhen 1 Hao	130	80	21 818	448	20.7	124	51.4	90.8	28.9	119.7	3.15
Mean		119	71	18 637	420	22.8	114	55.1	92.3	28.9	121.4	3.19
Standard deviation(s)		13.9	11.9	3 753.2	84.6	2.2	23.9	3.3	10.9	2.8	13.6	0.1

Continued next page.

Table 3. Continued.

		Stem trait					Seed trait								
Code	Name	Vegetative plant height	Vegetative node number	Vegetative internode length	Internode length	Stem diameter	Seeds per pod	100-seed weight	Seed protein content	Seed oil content	Palmitic acid	Stearic acid	Oleic acid	Linoleic acid	Linolenic acid
		mm	no.	mm	mm	mm	no.	g	g kg^{-1}	g kg^{-1}	mole %	mole %	mole %	mole %	mole %
1	Braxton	876	9.8	89	94	6.0	1.86	23.2	420	203	12.7	3.7	19.9	54.0	9.7
2	BSR 101	933	10.5	88	104	7.9	2.12	22.8	402	203	11.2	3.8	18.5	56.1	10.5
3	Century	676	9.8	69	97	7.9	2.06	21.6	438	192	11.8	3.4	17.2	57.3	10.3
4	Dassell	659	10	65	89	7.3	1.90	20.9	420	209	11.5	3.3	19.0	57.7	8.7
5	Dillon	825	10.8	76	93	8.2	1.92	21.8	419	208	11.1	2.3	27.4	50.6	8.6
6	Glenwood	830	10.4	79	105	8.1	2.05	20.3	412	205	11.6	3.8	18.6	56.5	9.6
7	Hutcheson	774	11.1	69	77	8.7	2.07	21.4	412	212	12.1	3.3	19.2	56.0	9.6
8	Johnston	829	10.8	76	91	6.8	1.77	22.4	429	201	12.6	3.1	24.8	51.4	8.1
9	KS 4694	763	10.5	72	88	8.7	2.08	23.8	415	205	12.1	3.1	23.5	53.0	8.4
10	Lambert	937	10.4	90	112	8.0	1.95	19.0	394	225	11.9	4.0	18.2	57.2	8.8
11	Macon	734	9.9	74	88	7.6	1.99	25.6	414	207	12.2	3.4	20.9	54.3	9.3
12	Manokin	677	10.4	65	80	7.1	2.00	17.1	418	206	12.4	3.7	21.3	54.2	8.5
13	Maple Donavan	749	10.8	69	84	7.8	2.09	19.9	409	211	12.2	3.3	19.6	55.8	9.2
14	McCall	730	10.5	69	93	7.6	2.17	16.4	419	199	12.2	3.8	16.5	57.4	10.1
15	Morgan	694	9.9	70	103	7.3	2.22	21.4	439	194	12.4	3.5	18.2	57.1	8.9
16	Ozzie	789	9.6	82	96	8.5	1.92	18.8	418	202	11.9	3.2	20.3	55.2	9.4
17	Parker	988	10.7	92	123	7.5	1.89	20.9	414	209	12.6	3.5	21.1	54.5	8.4
18	Pella	722	9.8	73	106	7.4	2.19	25.8	416	209	11.1	3.7	22.3	55.2	7.7
19	Preston	855	10.7	79	116	8.2	2.12	24.3	425	204	12.8	3.5	20.4	55.1	8.1
20	Sibley	661	9.8	67	93	7.9	1.88	20.6	415	209	12.4	2.9	22.5	53.7	8.7
21	Stonewall	838	10.1	83	106	6.9	2.29	23.7	417	212	12.2	3.3	22.8	54.2	7.7
22	Sturdy	738	9.7	76	95	7.1	2.31	21.9	421	203	11.5	3.2	22.2	54.1	9.1
23	TN 5-85	812	10.3	78	106	7.2	1.93	18.4	427	209	11.2	3.1	25.5	51.9	8.3
24	Weber	679	10	67	89	7.0	2.12	15.4	408	204	11.7	3.7	15.9	57.0	11.8
25	Young	781	11.3	69	89	8.3	1.99	22.5	442	197	12.6	2.5	20.8	55.3	8.8
26	De Dou 1 Hao	743	10.8	68	73	7.5	1.84	24.9	421	205	12.0	3.0	28.2	47.5	9.3
27	Jiu Nong 13	708	10.1	70	88	7.6	2.01	18.2	403	204	12.9	3.5	21.8	51.8	10.0
28	Dan Dou 5 Hao	783	10.8	72	91	8.4	1.75	25.0	446	200	11.5	3.3	20.1	55.8	9.5
29	Dong Nong 37	1086	11.5	94	98	8.4	2.07	23.7	422	196	12.9	3.9	17.6	53.9	11.7
30	Tong Nong 9 Hao	652	9.6	67	85	6.8	2.19	24.2	441	205	12.0	3.2	31.5	45.0	8.3
31	Nen Feng 9 Hao	736	10.1	72	82	8.7	2.15	23.6	433	204	11.7	3.5	25.5	51.1	8.2
32	Nen Feng 10 Hao	768	10.6	72	87	9.4	1.97	23.8	436	206	11.6	3.5	28.1	49.1	7.8
33	Hei Nong 29	1081	10.9	99	114	7.1	1.73	24.9	411	210	12.2	3.3	23.7	53.1	7.8
34	Ji Lin 18	729	11.1	65	80	6.9	2.19	22.6	456	189	11.5	2.8	32.8	45.8	7.1
35	Ji Lin 20	1006	10.6	95	122	7.7	1.86	26.2	407	205	12.1	3.1	36.5	40.8	7.0
36	Ji Lin 21	650	10.4	62	84	8.5	2.03	26.3	434	205	11.4	3.5	28.9	48.7	7.5
37	Kai Yu 8 Hao	815	10.9	74	84	7.9	1.82	26.5	421	205	11.6	2.9	33.1	44.4	7.2
38	Liao Nong 2 Hao	707	10.8	65	72	6.8	2.17	17.0	405	202	12.7	3.7	17.6	55.4	10.7
39	Lu Dou 4 Hao	1053	11	95	117	8.5	1.75	18.5	455	176	11.8	3.0	29.4	47.4	8.6
40	Lu Dou 7 Hao	899	10.4	86	106	7.5	1.81	19.4	447	183	12.4	3.3	20.9	54.9	8.6
41	Huai Dou 1 Hao	701	10.7	65	76	7.6	1.74	22.0	455	169	12.0	2.9	16.9	57.8	9.8
42	Hong Feng 3 Hao	894	10.7	83	89	6.6	1.96	18.5	425	193	13.2	3.7	18.7	52.7	11.8
43	Jiu Feng 1 Hao	994	10.4	95	107	6.5	1.76	20.0	432	207	11.7	3.2	20.1	54.4	10.7
44	Jiu Feng 2 Hao	986	10.8	91	95	5.8	2.20	17.4	434	188	12.0	3.8	16.5	57.1	10.6
45	He Feng 33	611	10	61	82	8.3	2.27	20.9	453	193	13.8	4.3	17.2	54.0	10.8
46	Hei Nong 37	1044	10.5	99	109	7.5	1.91	27.0	407	210	11.6	3.7	23.6	52.6	8.5
47	Ken Nong 2 Hao	797	9.9	80	84	7.2	2.14	22.0	402	205	14.1	3.6	19.9	53.0	9.4
48	Bai Nong 1 Hao	952	10.7	89	101	6.5	1.71	19.9	424	198	12.5	3.6	20.9	53.4	9.6
49	Tong Nong 8 Hao	843	10.1	83	97	6.8	1.77	23.0	431	203	11.8	3.1	23.1	53.3	8.7
50	Chen Dou 4 Hao	863	11.4	75	78	9.0	2.09	17.4	449	174	10.7	2.6	30.5	47.1	9.3
51	Gong Dou 4 Hao	914	10.1	90	99	6.1	1.70	26.3	432	185	11.3	2.9	33.2	45.2	7.5
52	Guan Dou 1 Hao	734	10.2	72	86	8.6	1.64	24.7	469	170	12.6	2.7	27.6	47.1	10.1
53	Zao Shu 9 Hao	779	10.4	74	89	6.9	1.77	24.1	453	186	12.1	2.4	29.6	47.0	8.3
54	Zao Shu 14	644	9.7	66	84	7.0	2.14	21.6	414	200	13.0	3.2	18.6	54.8	10.5
55	Zhe Chun 2 Hao	994	11.3	88	101	6.6	1.94	15.1	460	169	13.2	3.3	22.5	50.5	10.5
56	Ji Dou 7 Hao	625	9.6	65	91	9.5	2.22	28.2	442	200	11.5	2.7	34.1	44.6	7.2
57	Jin Dou 14	684	9.7	70	101	7.7	1.64	19.2	496	151	12.7	3.0	18.8	55.9	9.7
58	Jin Dou 15	720	10.2	70	78	6.8	1.83	23.2	437	200	12.7	2.7	19.3	57.0	8.4
59	Jin Dou 16	782	10.1	77	101	8.2	1.72	25.0	418	198	13.2	3.3	17.7	55.8	10.0
60	Liao Dou 10 Hao	573	9.8	58	73	8.8	1.57	31.2	453	190	11.7	2.9	50.2	29.8	7.9
61	Liao Dou 10 Hao	573	10.3	56	71	8.7	1.96	19.8	444	180	13.2	2.9	20.7	52.5	10.8
62	Tie Feng 22	778	10.6	73	92	8.0	2.25	26.6	442	210	11.1	2.6	41.4	38.7	6.2
63	Zhong Huang 1 Hao	477	9.3	51	66	8.8	1.97	21.1	433	192	12.3	3.1	19.8	54.5	9.5
64	Ji Dou 4 Hao	747	10.1	74	98	7.3	2.45	25.9	434	203	12.1	3.7	21.9	54.4	7.9
65	Jin Yi 9 Hao	714	10.5	68	96	7.7	1.94	26.4	434	200	11.7	2.6	32.0	47.4	6.2
66	Jin Yi 10 Hao	740	10.9	67	91	7.2	1.79	30.2	434	205	11.9	3.5	25.4	51.3	8.0
67	Yu Dou 8 Hao	753	10.3	73	86	7.2	2.17	23.2	449	198	12.1	3.4	22.7	52.5	9.3
68	Yu Dou 11	833	10.2	81	98	9.4	1.65	28.6	425	198	11.6	2.8	22.7	55.1	8.0
69	Zheng 133	590	9.7	60	82	9.3	1.96	25.4	457	187	11.4	2.7	19.8	56.1	10.0
70	Zheng 77249	906	9.8	92	102	7.5	1.96	24.8	440	186	12.5	2.6	22.9	52.9	9.1
71	Zhong Dou 19	751	10.2	73	84	7.8	1.75	21.3	434	165	12.4	2.9	21.9	53.1	9.9
72	Nin Zhen 1 Hao	843	10.9	77	98	8.7	2.22	22.1	464	164	12.2	2.9	28.5	47.1	9.4
Mean		789	10.4	76	93	7.7	1.97	22.4	430	197	12.1	3.2	23.5	52.2	9.0
Standard deviation(s)†		128.7	0.5	10.9	12.3	0.8	0.2	3.4	18.6	13.5	0.7	0.4	6.2	5.0	1.2

† Standard deviation of a genotypic mean.

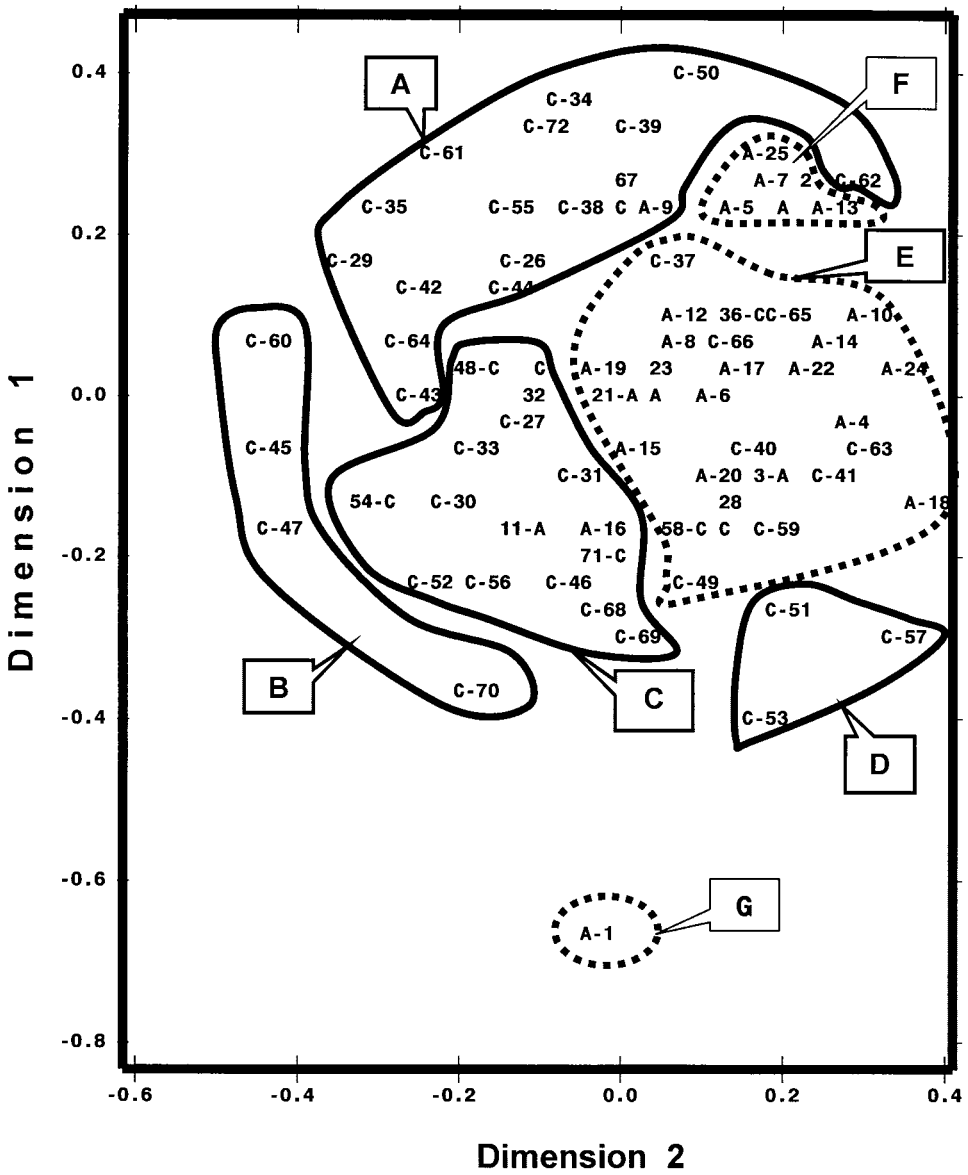

Fig. 1. Two-dimensional representation of genetic relationships among 72 modern Chinese and North American (NA) soybean cultivars derived from a two-dimensional multidimensional scaling (MDS) analysis determined on the basis of phenotypic similarity (PS) estimates. The PS was derived from multivariate analysis of phenotypic traits for plants grown in temperature- and photoperiod-controlled growth chambers. The stress value for the two-dimensional MDS analysis was 0.24 and the regression R^2 of fitted PS on the original PS was 0.74. A = NA cultivar, C = Chinese cultivar, numbers next to A or C are entry codes from Table 1. Seven clusters from Ward's minimum variance cluster analysis are superimposed on the MDS plot.

other methods available in the CLUSTER procedure. Ward's minimum variance method explained the greatest proportion of the variation in the PS matrix and was the only method from the CLUSTER procedure retained for further use. Because a nonhierarchical clustering approach had been shown to be effective in the cluster analysis of pedigree data, we also applied a nonhierarchical analysis to the present data set and compared results with those from Ward's minimum variance method (Cui et al., 2000b). In this separate alternative approach to clustering, the PS matrix was converted into Euclidean coordinates (71 dimensions) via the MDS procedure (SAS Institute, 1992). The MDS-derived coordinates were used subsequently as source data in a series of FASTCLUS analyses, where the analysis was optimized sequentially for 2 to 25 clusters. The FASTCLUS analyses employing three and seven clusters accounted for only 15 and 39% of the variation in the original PS matrix, whereas Ward's minimum variance method accounted for 42 and 59% for the same number of clusters. Because the MDS plus FASTCLUS approach was less effective than Ward's minimum variance method, only the output from Ward's minimum variance method was retained for analysis and interpretation.

Ward's minimum variance method produced a dendrogram of cultivar relationships (Fig. 2). The cubic clustering criterion and the pseudo *F* statistic had peaks at three clusters, indicating the existence of at least three major clusters (SAS Institute, 1985b; Milligan and Cooper, 1985). Examination of the dendrogram in rela-

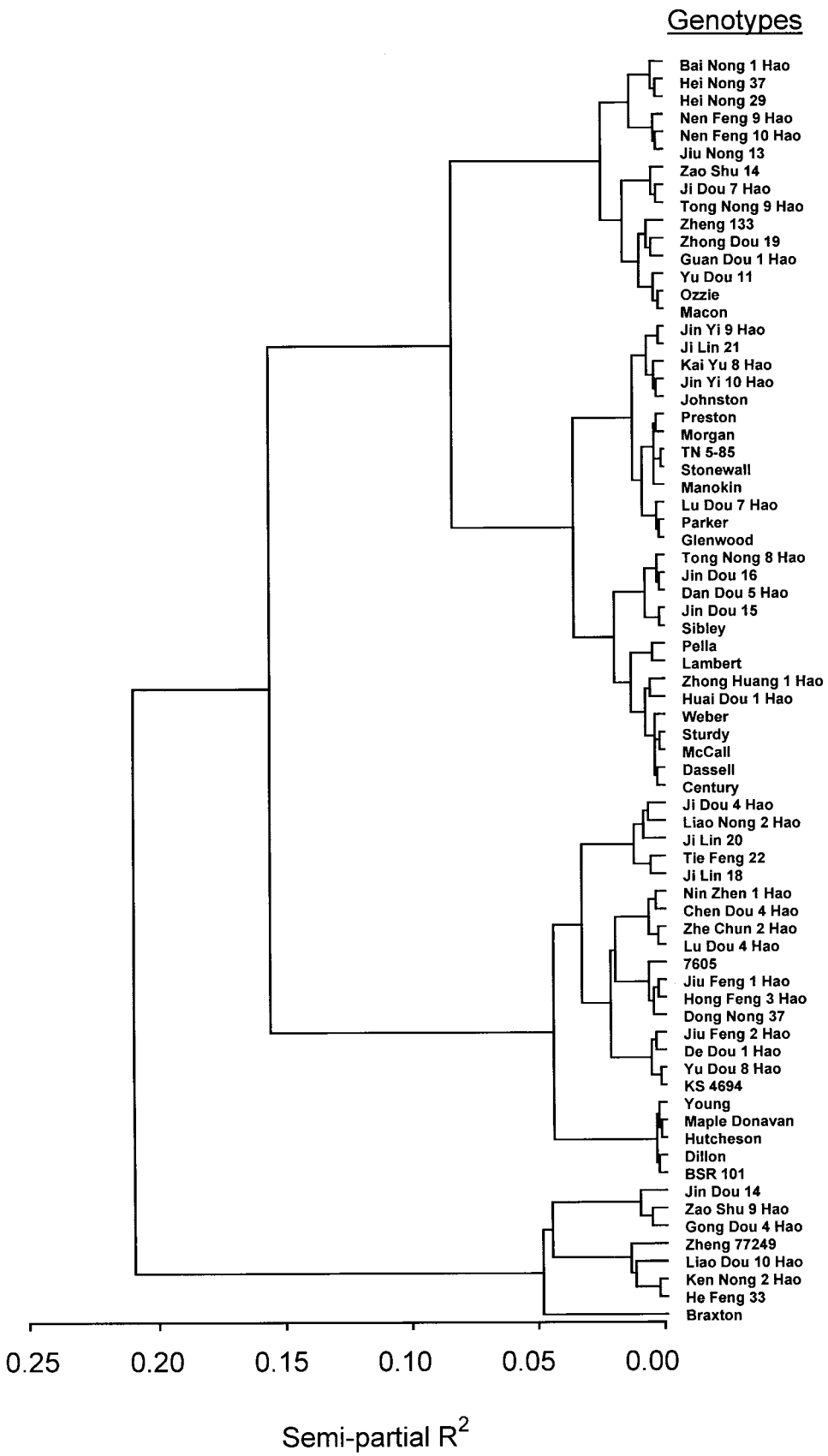

Fig. 2. Dendrogram of 47 Chinese and 25 North American modern soybean cultivars derived from phenotypic similarity estimates calculated from the first seven principal components by Ward's minimum variance cluster analysis.

tion to the origin and pedigree of cultivars revealed that each of the three major clusters could be decomposed further in a meaningful way to produce a total of seven clusters (Table 1 and Fig. 1). The decomposition of the three initial clusters into seven produced an increase in R^2 from 0.42 to 0.59.

Clusters A, B, C, and D were composed almost entirely of Chinese cultivars while Clusters F and G consisted only of NA cultivars. Cluster E was composed of both Chinese and NA cultivars. The seven clusters derived from Ward's minimum variance procedure corresponded well with 2-dimensional MDS-derived plots of PS (Fig. 1 and Tables 1, 4). Clusters A, B, and C were found on the left side in the figure and were early in maturity in terms of mean NA maturity groupings (between I and II), while Clusters D, E, F, and G on the right side of the figure were later in maturity (between II and VII). However, early maturity NA cultivars tended to separate from early maturity Chinese cultivars.

The 16 Chinese cultivars of Cluster A were geographically diverse in origin (i.e., from all three growing regions of China). These cultivars were tall during vegetative growth stages, and exhibited narrow leaves, short petioles, a high number of seeds per pod, small seed, and high linolenic acid content in the oil (Table 4). The four Chinese cultivars of Cluster B were from northern China and characterized as short plants with long narrow leaves, high leaf nitrogen content, low chlorophyll content, large seed, and high palmitic, stearic, and oleic, and lower linoleic acid content in the oil. The 13 Chinese cultivars of Cluster C were almost all from northern China and had thick leaves and long petioles. The three Chinese cultivars in Cluster D were from north central and southern China and had small round leaves, low leaf nitrogen content, low numbers of seed per pod, low oil and high protein content in the seed, and low stearic and linolenic acid content in the oil.

Cluster E was composed of 16 NA and 11 Chinese cultivars and exhibited intermediate values for most traits. Almost all Chinese cultivars in this cluster were from northern China and the NA cultivars in this cluster were derived largely from landraces from northern China. Although they appeared similar to each other phenotypically, no pedigree relation was found between these Chinese and NA cultivars. Cluster F contained five NA cultivars which exhibited thin leaves, large stem diameter, high oil, linoleic acid, and chlorophyll contents, and low seed protein content. Cluster G had only one member, the NA cultivar Braxton, which was unique phenotypically and represented the genetic extreme for several traits in this study (Table 4). The pedigree of Braxton was also unique in that 19% of its ancestry traced to the

Table 4. Genotypic composition and trait means for seven clusters formed from 72 Chinese and North American (NA) soybean cultivars based on phenotypic similarity (PS). The PS estimates were derived from multivariate analysis of phenotypic traits for plants grown in the temperature- and photoperiod-controlled growth chambers.

	PS-based cluster						
	Predominantly Chinese Clusters				Mixture of Chinese and NA	Predominantly NA clusters	
Characterization	A	B	C	D	E	F	G
Genotypic composition							
Average U.S. maturity group†	1.5	1.3	1.7	2.7	2.6	3.6	7.0
Chinese cultivars, No.	16	4	13	3	11	0	0
NA cultivars, No.	1	0	2	0	16	5	1
Total cultivars, No.	17	4	15	3	27	5	1
Leaf trait							
Leaf length, mm	126.5	131.4	128.6	95.5	112.1	114.5	77.4
Leaf width, mm	62.8	69.5	75.5	63.7	75.0	73.7	50.6
Area per leaf, mm^2	17 599.8	19 793.3	21 139.8	13 429.3	18 708.1	18 535.2	8 303.0
Dry weight per leaf, mg	390.9	466.6	496.5	292.1	413.2	391.1	321.2
Specific leaf weight, g m^{-1}	22.8	23.6	23.6	22.9	22.1	21.2	32.5
Petiole length, mm	106.0	114.4	130.0	113.8	110.8	123.6	73.5
Leaf nitrogen content, g kg^{-1}	54.5	59.0	54.7	49.4	56.0	54.9	48.2
Chlorophyll a content, g kg^{-1}	93.3	85.0	90.3	87.3	93.0	100.3	91.2
Chlorophyll b content, g kg^{-1}	29.4	27.2	28.3	27.8	29.0	31.4	26.8
Chlorophyll a+b content, g kg^{-1}	122.7	112.2	118.9	115.0	122.2	131.7	118.0
Chlorophyll ratio, ratio	3.2	3.1	3.2	3.2	3.2	3.2	3.4
Stem trait							
Vegetative plant height, mm	854.0	721.8	776.1	792.3	757.4	812.4	876.0
Vegetative node number, No.	10.8	9.9	10.1	10.1	10.3	10.9	9.8
Vegetative internode length, mm	79.2	73.1	76.5	78.6	73.6	74.7	89.4
Internode length, mm	92.0	85.2	91.6	96.5	95.3	89.5	94.2
Stem diameter, mm	7.6	7.9	8.1	6.9	7.6	8.2	6.0
Seed trait							
Seeds per pod, No.	2.1	2.0	1.9	1.7	2.0	2.0	1.9
100-Seed weight, g	21.3	24.7	23.7	23.2	22.2	21.7	23.2
Seed protein content, g kg^{-1}	436.0	436.7	428.4	460.3	425.9	416.8	419.5
Seed oil content, g kg^{-1}	191.9	193.6	197.6	174.0	201.7	206.1	202.8
Palmitic acid content, mole %	12.1	13.0	12.1	12.0	12.0	11.8	12.7
Stearic acid content, mole %	3.2	3.4	3.2	2.8	3.3	3.0	3.7
Oleic acid content, mole %	25.2	27.5	24.1	27.2	21.6	21.1	19.9
Linoleic acid content, mole %	50.0	47.4	51.7	49.3	54.3	54.7	54.0
Linolenic acid content, mole %	9.4	9.3	9.0	8.5	8.7	9.3	9.7

† For ease of calculation and representation of means, maturity group data were converted to Arabic rather than standard Roman numerals, where 000 = −2, 00 = −1, 0 = 0, I = 1, II = 2, III = 3, etc.

Japanese cultivar Tokyo. No other NA entry in the study had as high a percentage of Japanese germplasm in the pedigree.

Comparison of PS-Based and CP-Based Clusters

Because clear PS-based clusters were identified in this study, we sought to validate their integrity through a comparison with CP-based clusters described previously by Gizlice et al. (1996) (North America) and Cui et al. (2000b) (China) (Table 1). Comparisons were made in the following way. All cultivars in this study were assigned to a PS-based cluster. Many were also assigned to CP-based clusters in previous studies. Discrepancies between the former and latter assignments provided the basis for comparison of PS and CP results. A total of 20 Chinese and 16 NA cultivars were available to compare CP and PS cluster assignments (Table 1). Results showed a trend that members from a CP-based cluster were usually assigned to a single PS-based cluster. Thus, a positive relation existed between CP- and PS-based clusters. For example, the eight pairs of NA cultivars representing eight CP-based clusters were assigned to only two PS-based clusters. For six of the eight pairs, members of a CP based cluster were assigned to the same PS-based cluster. Groups of Chinese cultivars representing eight CP-based clusters were assigned to four PS-based clusters. For 12 of the 20 Chinese cultivars, all members of a CP-based cluster were also assigned to a single PS-based cluster. Where PS-based cluster assignments differed from CP-based assignments, the cultivars were often near each other on the graph of PS derived from MDS analysis (Fig. 1). This agreement tended to validate the PS-based clustering approach.

Relationship between PS and CP

China

Because PS- and CP-based clusters tended to show some agreement in characterization of cultivars within a country, we also examined the overall correlation of CP and PS estimates within a country. The mean CP relation for the 47 Chinese soybean cultivars was 0.03 and ranged from 0 to 0.63, with distribution of CP values strongly skewed toward values below 0.25 (Cui et al., 2000b). The mean PS relation for this group was higher (0.58) and ranged from 0.24 to 0.88. A weak but significant positive correlation ($r = 0.14$, $P < 0.01$) was found between the CP and corresponding PS values. The weakness of this linear association may have resulted partly from the preponderance of low CP values. Based on empirical results from Manjarrez-Sandoval et al. (1997), CP and PS would be expected to have little relation for CP values below 0.25 (the majority of CP values here), while the relation between genetic PS and CP should increase for higher CP values. When the cultivar pairs with low CP (<0.25) were dropped from analysis, the correlation between CP and corresponding PS rose from 0.14 to 0.43 ($P < 0.01$).

The correlation of PS with CP would also be expected to rise as the number of measured traits increases, because PS would then more completely represent the true underlying genetic pool effects. Had we employed a larger number of phenotypic traits in this study, the relation may have improved beyond the observed 0.43. Van Beuningen and Busch (1997) employed 35 traits and found a significant correlation of 0.68 between CP and PS.

North America

The mean and range in CP for 25 NA soybean cultivars was 0.145 (0.00–0.62) and the corresponding mean and range for PS was 0.69 (0.06–0.92). The correlation between CP and PS was low but significant and similar to that found for Chinese cultivars ($r = 0.13$, $P = 0.02$). The relation of PS and CP for the 25 NA soybean cultivars illustrated the same trend observed for Chinese cultivars. However, deletion of CP values below 0.25 did not improve the correlation between CP and PS, because few pairs had CP relationships above 0.25.

Five Chinese cultivars contained NA parentage in the pedigree (Table 1) and were a potential basis for comparing the impact of innate diversity (derived from founder effects) vs. breeder selection pressure on phenotype and PS estimates. If these five 'bridge cultivars' appeared more similar to Chinese than NA cultivars, phenotypically, then the result would suggest that contrasting breeder selection pressure on phenotype in China and North America may be important in the interpretation of our results. By contrast, if bridge cultivars appeared intermediate, phenotypically, between Chinese and U.S. cultivars, then this result would suggest that innate founder effects might be more important than breeder selection on PS estimates. Unfortunately, the parents (the appropriate controls for examining breeder selection effects) of the five bridge cultivars were not evaluated in this study; and, none of the other cultivars were closely related to the bridge cultivars. Thus, interpretation of the PS relationships between the bridge cultivars and Chinese or NA cultivars was unclear. However, it is interesting to note that none of the five bridge cultivars fell into clusters dominated by NA cultivars and that each was phenotypically more similar to Chinese than NA cultivars. Thus, there is a suggestion that breeder selection effects on PS may have been important.

Factors Affecting Genotypic Variation in PS

Maturity group, geographic origin, and seed yielding ability, as well as PS-based clustering of cultivars can all be important factors, which explain or classify phenotypic diversity in soybean breeding programs. We examined the association between these factors (Table 1) and PS by subjecting portions of the 72 by 72 PS matrix to single and multiple-factor regression analyses with factors, such as maturity group, treated as independent class variables (Table 5). For all factors except clustering results and maturity groups, Chinese and NA cultivars were examined in separate regression analyses. For Chinese cultivars, only PS-based clusters were associated with cultivar variation in the PS matrix. For NA cultivars, maturity group was associated with PS ($R^2 = 0.37$) and mirrored a relation between CP and maturity group

Table 5. Effectiveness of cluster analysis, maturity group (MG), province or region of origin, and yield performance in explaining variation in phenotypic similarity (PS) among Chinese and North American (NA) cultivars. Effectiveness was measured using coefficient of determination (R^2) and obtained from a series of single and two-factor regression analyses. Yield and maturity ratings for Chinese cultivars were collected from a minimum of five NA environments. Yield was expressed as a percentage of NA control cultivars for analysis. Phenotypic similarity estimates were employed as the dependent variable in the analyses. The PS estimates were derived from 25 metric traits recorded for Chinese and NA soybean cultivars grown in growth chambers.

	R^2					
	Chinese		North American		All	
Factor	Single factor analysis	Two factor analysis of cluster plus one additional factor‡	Single factor analysis	Two factor analysis of cluster plus one additional factor‡	Single factor analysis	Two factor analysis of cluster plus one additional factor‡
Cluster	0.41	–	0.83	–	0.59	–
MG	0.06	0.47	0.37	0.88	0.10	0.62
Region†	0.03	0.44	0.08	0.83	–	–
Province	0.14	0.50	–	–	–	–
Yield level	0.03	0.44	–	–	–	–

† Regions of China: northeastern China, northern China, and southern China; regions of North America: northern part of North American and Southern United States.
‡ The second factor in the analysis is identified in the far left column.

detected previously by Gizlice et al. (1996) (Table 5). This relationship between PS and maturity group was related to a strong NA breeder. However, multiple regression analysis indicated that maturity effects were correlated almost completely with cluster effects. In summary, the seven PS-based clusters explained 59% of the variation in PS, indicating that cluster analysis was clearly the best discriminator of phenotypic diversity in this study. Inclusion of multiple factors in the regression analysis did not improve the R^2 appreciably over that obtained by cluster analysis alone.

Maturity difference between cultivars was a better indicator of phenotypic diversity for NA than Chinese cultivars, because maturity was a good indicator of the geographical origin of a cultivar in North America. In China, however, cropping systems for soybean are such that a group II, III, or IV cultivar could have arisen from either the South or the North, so that one cannot predict geographical origin well by maturity alone. Cui et al. (2000b) showed that CP patterns were in part a function of geographical origin.

Implications to Soybean Breeding

The purpose of this study was to examine the hypothesis that NA and Chinese cultivars shared a common, though unrecorded, underlying genetic base. Experiments revealed that phenotypic diversity was much greater in Chinese than in NA cultivars (Fig. 1, Fig. 2, and Table 2). The results suggest that, although most NA and Chinese cultivars trace their pedigree to Chinese landraces, NA cultivars may be derived from a subset of the Chinese cultivar genetic base. In suppporting research, Li et al. (2001) compared Chinese and NA soybean ancestors using RAPD markers and found that some ancestors of NA breeding programs were similar genetically to ancestors used in Chinese breeding. It is also possible that decades of breeder selection for adaptation to contrasting environmental conditions in China and North America may have accentuated the genotypic and, thus, the phenotypic diversity between Chinese and NA cultivars. The findings here support the concept that Chinese and NA cultivars are phenotypically distinct and, thus, potentially good genetic sources for broadening the agronomic, morphological, and biochemical diversity of the contrasting breeding programs.

Neither the trait-based PS analysis reported here nor previously reported CP analysis were perfect measures of diversity, because they both depended on assumptions that may or may not have been met completely. However, they are sufficiently useful such that breeders may consider the use of PS and CP as well as agronomic performance as guides in the selection of exotic parents for local breeding (Cox et al., 1985; van Beuningen and Busch, 1997). We speculate that specific hybrid combinations between the two genotypic pools may be chosen best by avoiding matings within phenotypic-based PS clusters and by selecting parents with high yield. Agronomic studies have shown that 10 of the 47 Chinese cultivars studied here are relatively high yielding in the North America (Table 1). Only two of the 10 were associated with phenotypic clusters dominated by NA cultivars. If high protein content is a priority for a breeder, it was noted that cultivars from central and south China tended to have high protein content in the seed, especially those in the PS-based Cluster D.

China has successfully used elite NA cultivars as breeding stock, demonstrating the effectiveness of this approach in broadening the genetic base for future breeding. Contrary to the conventional expectations regarding the difficulty in using exotic materials in applied breeding, most new Chinese cultivars derived from U.S. cultivars yielded well. For example, 'Si Dou 11' was selected from the cross of 'Si Dou 2 Hao' × 'Williams' and released in 1987 in Jiangsu province. It has ranked high in Chinese yield trials to the present. 'Ji Dou 7 Hao' was developed from the cross Williams × 'Cheng Dou 1 Hao' and released in 1992 in Hebei province. It established a yield record of 4749 kg ha^{-1} in provincial yield trials (Cui et al., 1998).

It is interesting to note that among the NA cultivars studied here, both maturity group and PS-based cluster were useful in explaining phenotypic diversity. To maximize morphological diversity (and presumably agronomic diversity) in the progeny from NA × NA culti-

vars examined here, one might want to choose parents that differ in maturity, as well as PS and CP based cluster designation.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Judith F. Thomas and her staff of the Southeastern Plant Environment Laboratories at North Carolina State University for their assistance in use of the phytotron facilities.

REFERENCES

Autrique, E., M.M. Nachit, P. Monneveux, S.D. Tanksley, and M.E. Sorrells. 1996. Genetic diversity in durum wheat based on RFLPs, morphophysiological traits, and coefficient of parentage. Crop Sci. 36:735–742.

Carter, T.E., Jr., Z. Gizlice, and J.W. Burton. 1993. Coefficient-of-parentage and genetic similarity estimates for 258 North American soybean cultivars released by public agencies during 1945–1988. U.S. Dep. of Agric. Tech. Bull. Num. 1814. U.S. Govt. Print. Office, Washington, DC.

Carter, T.E., Jr., R.L. Nelson, P.B. Cregan, H.R. Boerma, P. Manjarrez-Sandoval, X. Zhou, W.H. Kenworthy, and G.N. Ude. 2000. Project SAVE (Soybean Asian Variety Evaluation)—potential new sources of yield genes with no strings from USB, public, and private cooperative research. p. 68–83. *In* B. Park (ed.) Proceedings of the twenty-eighth soybean seed research conference 1998. American Seed Trade Association, Inc., Washington, DC.

Carver, B.F., R.F. Wilson, and J.W. Burton. 1984. Developmental changes in acyl-composition of soybean seed selected for high oleic acid concentration. Crop Sci. 24:1016–1019.

Chang, R., L. Qiu, J. Sun, Y. Chen, X. Li, and Z. Xu. 1999. Collection and conservation of soybean Germplasm in China. p. 172–176. *In* H.E. Kauffman (ed.) Proc. World Soybean Research Conference VI, Chicago, IL. 4–7 Aug. 1999. Superior Printing, Champain, IL.

Chang, R., J. Sun, and L. Qiu. 1992. Introduction and utilization of exotic soybean germplasm. *In* Shenglin Li et al. (ed.) Plant specific gene resources and their utilization. p. 114–116. China Science and Technology Press, Beijing.

Cox, T.S., Y.T. Kiang, M.B. Gorman, and D.M. Rodgers. 1985. Relationship between coefficient of parentage and genetic similarity in the soybean. Crop Sci. 25:529–532.

Cui, Z., T.E. Carter, Jr., and J.W. Burton. 2000a. Genetic base of 651 Chinese soybean cultivars released during 1923 to 1995. Crop Sci. 40:1470–1481.

Cui, Z., T.E. Carter, Jr., and J.W. Burton. 2000b. Genetic diversity patterns in Chinese soybean cultivars based on coefficient of parentage. Crop Sci. 40:1780–1793.

Cui, Z., T.E. Carter, Jr., J. Gai, J. Qiu, and R.L. Nelson. 1999. Origin, description and pedigree of Chinese soybean cultivars released from 1923 to 1995. U.S. Dep. of Agric. Tech. Bull. 1871. U.S. Govt. Print. Office, Washington, DC.

Cui, Z., J. Gai, T.E. Carter, Jr., J. Qiu, and T. Zhao. 1998. The released Chinese soybean cultivars and their pedigree analyses (1923–1995). China Agriculture Press, Beijing.

Downs, R.J., and V.P. Bonaminio. 1976. Phytotron procedural manual for controlled environmental research at the Southeastern Plant Enviroment Laboratories. North Carolina Agric. Exp. Stn. Tech. Bull. No. 244.

Fehr, W.F. 1999. The importance and future prospects for genetic diversity for yield improvement. p. 186. *In* H.E. Kauffman (ed.) Proc. World Soybean Research Conf. VI. 4–7 Aug. 1999. Superior Printing, Champaign, IL.

Gai, J. 1997. Plant breeding: Crop species. (In Chinese). China Agriculture Press, Beijing.

Gizlice, Z., T.E. Carter, Jr., and J.W. Burton. 1993. Genetic diversity in North American soybean: I. Multivariate analysis of founding stock and relation to coefficient of parentage. Crop Sci. 33:614–620.

Gizlice, Z., T.E. Carter, Jr., and J.W. Burton. 1994. Genetic base for North American public soybean cultivars released between 1947 and 1988. Crop Sci. 34:1143–1151.

Gizlice, Z., T.E. Carter, Jr., T.M. Gerig, and J.W. Burton. 1996. Genetic diversity patterns in North American public soybean cultivars based on coefficient of parentage. Crop Sci. 36:753–765.

Glowa, W. 1974. Zirconium dioxide: a new catalyst in the Kjeldahl method fortotal nitrogen determination. J. Assoc. Off. Anal. Chem. 57:1228–1230.

Goodman, M.M. 1972. Distance analysis in biology. Syst. Zool. 21: 174–186.

Grafius, J.E. 1978. Multiple characters and correlated response. Crop Sci. 18:931–934.

Grafius, J.E., R.L. Thomas, and J. Bernard. 1976. Effect of parental component complementation on yield and components of yield in Barley. Crop Sci. 16:673–677.

Johns, M.A., P.W. Skroch, J. Nienhuis, P. Hinrichsen, G. Bascur, and C. Munoz-Schick. 1997. Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. Crop Sci. 37:605–613.

Li, Z., L. Qiu, J.A. Thompson, M.M. Welsh, and R.L. Nelson. 2001. Molecular genetic analysis of U.S. and Chinese soybean ancestral lines. Crop Sci. 41:1330–1336.

Manjarrez-Sandoval, P., T.E. Carter, Jr., D.M. Webb, and J.W. Burton. 1997. RFLP genetic similarity estimates and coefficient of parentage as genetic variance predictors for soybean yield. Crop Sci. 37:698–703.

Martin, B.A., R.F. Wilson, and R.W. Rinne. 1986. Temperature effects on the expression of a high-oleic acid trait in soybean. J. Am. Oil Chem Soc. 63:346–352.

Milligan, G.W., and M.C. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50:159–179.

Moran, R. 1982. Formulae for determination of chlorophyllous pigments extracted with *N*, *N*-dimethylformamide. Plant Physiol. 69: 1376–1381.

Moran, R., and D. Porath. 1980. Chlorophyll determination in intact tissues using *N*, *N*-dimethylformamide. Plant Physiol. 65:478–479.

Nelson, D.W., and L.E. Sommers. 1973. Determination of N in plant material. Agron. J. 65:109–112.

SAS Institute. 1985a. SAS user's guide: Basics. 5th ed. SAS Institute Inc., Cary, NC.

SAS Institute. 1985b. SAS user's guide: Statistics. 5th ed. SAS Institute Inc., Cary, NC.

SAS Institute. 1992. Technical report P-229. SAS Institute Inc., Cary, NC.

Sneller, C.H. 1994. Pedigree analysis of elite soybean lines. Crop Sci. 34:1515–1522.

Sneller, C.H. 1999. Diversity within commercially used gene pools. p. 176–184. *In* H.E. Kauffman (ed.) Proc. World Soybean Research Conf. VI. Aug. 4–7, 1999. Superior Printing, Champaign, IL.

Specht, J.E., D.J. Hume, and S.V. Kumudini. 1999. Soybean yield potential—A genetic and physiological perspective. Crop Sci. 39: 1560–1570.

Specht, J.E., and J.H. Williams. 1984. Contribution of genetic technology to soybean productivity—retrospect and prospect. p. 49–74. *In* W.R. Fehr (ed.) Genetic contributions to yield gains of five major crop plants. Agron. Monogr. 7. CSSA, Madison, WI.

Thompson, J.A., L. Qiu, R.L. Nelson, and Z. Li. 1997. Genetic diversity of NA and Chinese ancestral soybean lines. p. 160. *In* Agronomy Abstracts. ASA, Madison, WI.

van Beuningen, L.T., and R.H. Busch. 1997. Genetic diversity among North American spring wheat cultivars: I. Analysis of the coefficient of parentage matrix. Crop Sci. 37:570–579.

Zhou, X., T.E. Carter, Jr., R.L. Nelson, H.R. Boerma, and B.R. McCollum. 1998. Breeding progress vs. breeding effort: the road ahead in soybean variety development. p. 81. *In* Agronomy Abstracts. ASA, Madison, WI.